Cost-Effectiveness Analysis in Radiology

In the past 20 years, there has been increasing recognition of the need to consider cost in medical decision making. This period has seen an explosion in the number of economic evaluations appearing in the medical literature. Cost-effectiveness analysis is an objective systematic technique for comparing alternative health care strategies on both cost and effectiveness simultaneously. Cost-effectiveness analysis can be used to inform medical decision makers in the establishment of clinical practice guidelines and in the setting of health policy. Cost-effectiveness analysis is a state-of-the-art research tool with its own terminology and methods. It is critical that radiologists become familiar with the concepts and procedures of cost-effectiveness analysis so they can properly evaluate cost-effectiveness analysis studies and be more knowledgeable participants in the health care decision-making process. This article explains the rationale, terminology, and methods of cost-effectiveness analysis as applied to radiology.

In radiology, as in all of medicine, challenging decisions must be made every day. Without systematic analysis, it is difficult to identify clearly the relevant alternatives, assimilate the vast amounts of relevant data present in practice and the literature, and then make sound judgments. In a world of limited health care resources, medical decision makers must consider the benefits of an intervention in relation to its cost. In the absence of explicit consideration of cost, it is inevitable that health care resources will be inefficiently allocated, which results in reduced overall health benefits for the population. Cost-effectiveness analysis (CEA) is one tool to support rational medical decision-making in a cost-conscious environment.

CEA is a research technique that has become increasingly important to the practice of medicine during the past 20 years. The number of imaging technology studies that focus on cost-effectiveness and cost-benefit analyses has nearly quadrupled during the same period (1). However, these terms are not always used correctly and are often misunderstood (2). When a diagnostic test is said to be “cost-effective,” the additional health benefits it provides are considered to be a good value for the money spent, as compared with those of the alternatives. CEA is an objective technique used to assess whether a new or more effective test or treatment is worth the additional cost. CEA is an extension of clinical decision analysis, an established methodology for evaluating complex medical decisions with uncertain outcomes.

CEA is used, either formally or informally, by hospitals, the federal government, managed care organizations, self-insured employers, researchers, and individual physicians. With increasing frequency, health policymakers use CEs when they decide whether to implement new interventions and health programs. CEs can also contribute to the development of clinical practice guidelines and to coverage and reimbursement decisions. As a general example, studies on the cost-effectiveness of the Pneumococcus vaccine in the elderly were instrumental in the decision by Medicare to reimburse for this vaccine (3,4). CEA is not intended to stand alone in the decision-making process but can provide an insightful systematic perspective. As compared with randomized controlled trials, CEA models have the advantage of being relatively inexpensive and can be used to test strategies that would be difficult to test in actual practice. CEA may also guide researchers in a new direction for further research. Finally, CEA may serve as justification for more definitive and more costly randomized trials.

There are clear signs that the CEA techniques are maturing. Several journals have published guidelines for CEs (5,6). The U.S. Public Health Service established a panel on cost-effectiveness in health and medicine to create detailed standards for CEs. In 1996,
the panel’s full report was published in book form (7), and summaries have appeared in journals (8–10). The report’s goals were to improve the quality, standards, and consistency of published CEs, which would allow fair comparisons.

This article reviews the unique vocabulary and methods of and rationale for the use of CEA in medicine, knowledge of which will enable the radiology community to participate more fully in setting health care policy. A summary checklist of the components of a cost-effectiveness model and a recommended reading list are included in Appendices 1 and 2, respectively.

**TYPES OF ECONOMIC ANALYSES IN HEALTH CARE**

A full economic evaluation considers both cost and effectiveness for the relevant set of choices. Partial evaluations examine cost alone, effectiveness alone, or a single diagnostic or treatment strategy. The terms “outcome” and “effectiveness” are often used interchangeably in the literature. The term “cost-effectiveness analysis” is commonly used in a general sense to refer to any of the four types of full economic evaluation (Table 1) (11).

**Cost-Minimization Analysis**

Cost-minimization analysis is a comparison of the costs of different health-care strategies that are assumed to have similar effectiveness. Although the assumption of equal effectiveness limits the use of this technique, cost-minimization analyses appear occasionally in the radiology literature (12–16).

**Cost-Benefit Analysis**

Cost-benefit analysis uses monetary units such as dollars to compare the costs of a health intervention with its health benefits. In essence, cost-benefit analysis converts all benefits to a cost equivalent. Cost-benefit analysis is a fundamental technique in the world of finance and economics but has had limited success in health care. If a new treatment costs more than the standard treatment but can be used to save lives, the life-years saved are converted into an equivalent dollar amount. This method has the advantage of valuing everything in the same units, which permits direct comparison of treatment costs with health benefits, as well as comparison of health programs with nonhealth programs. The difficulty in assigning a monetary value to a health scenario has limited its use and acceptance in the medical field (17,18).

**Cost-Effectiveness Analysis**

CEA refers to analyses that consider both cost and effectiveness, where effectiveness is an objective measure (eg, the number of cancers diagnosed or life-years saved). The authors of a landmark article (19), who used this technique in the medical field, evaluated the cumulative cost-effectiveness of serial stool guaiac tests for colon cancer screening, with cancers detected being the effectiveness measure. To the astonishment of policy experts, the additional cost of progressing from the fifth to the sixth stool guaiac test was $47 million per additional case detected. Objective measures of effectiveness may be preferred in situations in which the time horizon is short and the accompanying reduction in quality of life is unlikely to have much effect. Radiology CEs often use intermediate outcomes, such as the length of stay for a hospital admission, readmission rates, or the number of unnecessary surgeries prevented (20,21).

**Cost-Utility Analysis**

Cost-utility analysis is similar to CEA, except that the effectiveness measure is a subjective measure and explicitly incorporates patient preferences. In some cases, this may be as simple as recognizing that a patient would prefer to live in full health rather than have a chronic illness. In other situations, the issue may be more subtle. For example, a man who needs to undergo evaluation for possible peripheral vascular disease may choose to wait for a later appointment for magnetic resonance (MR) angiography rather than undergo conventional angiography, which has risks and discomfort but no waiting time for an appointment (22). Cost-utility analysis incorporates the patient’s subjective value of the risk, discomfort, and waiting time into the effectiveness measure of these two diagnostic examinations. Ultimately, all medical decisions are the patient’s and must reflect the values and priorities of the patient (23). This is why cost-utility analysis is the preferred method for conducting economic evaluations.

The most commonly used measure in cost-utility analysis is the quality-adjusted life-year (QALY). The basic concept behind the QALY is that 1 year of life with substantial morbidity is less desirable than 1 year of excellent health. The QALY incorporates this quality of life adjustment through the use of a preference weight for each state of health on a scale from zero to one, where zero represents death and one represents full health. This is often referred to as a utility score, hence the name “cost-utility analysis.” The utility score for each state of health is then multiplied by the length of time spent in that state of health. One year of full health counts as 1 QALY. A major illness with a utility score of 0.5 would require 2 years of life with this major illness to equal 1 QALY. To further illustrate this concept, let us assume that the utilities of mild, moderate, severe angina are 0.9, 0.7, and 0.5, respectively (24). A patient lives 17 years: 10 years with mild angina, 5 years with moderate angina, and 2 years with severe angina, and then dies. To calculate the number of QALYs for this patient, multiply the time spent in each state of health by its associated utility score and sum them. Thus, we have $10 \times 0.9 + (5 \times 0.7) + (2 \times 0.5) = 13.5$ QALYs. On the basis of this analysis, 14 years of excellent health would be preferable to the 17 years this patient had angina.

For simplicity, the term “CEA” is used in the remainder of this article to refer to any type of full economic evaluation.

**TABLE 1**

<table>
<thead>
<tr>
<th>Evaluation Method</th>
<th>Valuation of Effectiveness</th>
<th>Compares Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost-minimization analysis</td>
<td>Assumed equal dollars</td>
<td>Yes</td>
</tr>
<tr>
<td>Cost-benefit analysis</td>
<td>Dollars</td>
<td>Yes</td>
</tr>
<tr>
<td>CEA</td>
<td>Objective measure (eg, life-years saved)</td>
<td>Yes</td>
</tr>
<tr>
<td>Cost-utility analysis</td>
<td>Subjective measure (eg, life-years adjusted for quality of life)</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Note.—Data are adapted from Drummond et al (11). Valuation of cost was in dollars for all methods.
DEFINING THE CEA MODEL

As in any research, a cost-effectiveness study must define the problem it will address. The problem definition can be broken into five parts: reference case, strategies, time horizon, perspective, and effectiveness measure. These parts provide the context for the study. They also allow the reader to judge whether the problem being addressed is important, timely, and relevant.

Reference Case

The reference case defines a typical patient among the clinical population of interest. This is analogous to defining the study population in the methods section of a radiology manuscript. For example, a researcher studying the imaging diagnosis of stroke might use a reference case of a 65-year-old man presenting to the emergency department.

The usual approach is to perform a computer simulation of a cohort of patients that match the reference case and then repeat the analysis for each subgroup of interest. By using the stroke example, it may be useful to policymakers to examine a subgroup of younger men, black men, or women to determine the appropriateness of imaging in these populations. In another example, the accuracy of mammography differs across age groups and therefore may result in different conclusions for each subgroup (25).

One of the difficulties in defining the reference case is the limited availability of data that are appropriate for the specified reference case. This limitation often forces the choice of reference case to be based on the populations included in published studies. As an alternative, one can use expert opinion or generalize data from the studied population to the reference case in the model. However, these options should be exercised with caution because they increase the level of uncertainty in the model.

Strategies

CEAs generally address whether a particular strategy should become the standard of care. This requires careful selection of the appropriate set of alternative strategies (26). If important diagnostic or treatment options are excluded, then the model will be irrelevant. However, including many choices will not only complicate the analysis and the message but will also usually require more assumptions, which will weaken the analysis. Thus, analyses are frequently limited to two choices.

The selection of strategies must reflect actual practice. Clinical trials usually focus on the efficacy of specific diagnostic tests or treatments and address questions such as “Is imaging test A superior to imaging test B?”, where test B may be a standard imaging technology and test A a new technology. For example, Garcia Pena et al (27) used a decision model for children with abdominal pain at presentation to the emergency department to compare three strategies of computed tomography (CT) use with current practice. The three strategies were to (a) perform CT in all children and discharge those with normal studies, (b) perform CT in all children and admit all, and (c) selectively perform CT in those children who had a white blood cell count of at least 10 × 10^9/L and admit all. All three strategies reduced the cost of care, as compared with current practice. Einstein et al (28), who investigated subsequent imaging in patients who had renal masses at urography, used a variation on this approach: CT alone, ultrasonography (US) alone, CT followed by US, or US followed by CT. They concluded that the most cost-efficient work-up was US followed by CT, when necessary.

Time Horizon

A CEA must state the period for which benefits and risks are considered. In an ideal situation, the time horizon should be long enough to fully capture the benefits and risks of each strategy. However, the time horizon may be limited by the data available to support the model. Extrapolation of results of short-term studies to the long term inevitably adds uncertainty. Sometimes the model may be run for a short time, but the effectiveness measure may account for the improved life expectancy beyond the end of the model’s time horizon. For example, the benefit of a 5-year cancer detection program will come from the added years of life well beyond the 5-year time horizon.

Perspective

The perspective defines the viewpoint of the decision maker. Determining which perspective to adopt is a subtle yet crucial component of defining a cost-effectiveness problem. The societal perspective evaluates all health effects and changes in resource use. It has been recommended because it is the most comprehensive viewpoint and allows for easy comparison with other CEA studies, but it is often insightful to consider other perspectives as well (29). Such perspectives might include that of an individual, a managed care insurer, a single hospital, a health care system, Medicaid, Medicare, or the federal government.

Although effectiveness is typically measured in terms of the benefit to the patient, the choice of perspective determines which costs are included in the analysis. Patients who have full insurance experience no cost if they undergo CT. The insurer must consider not only the procedure reimbursement cost but also the cost of processing the claim. If the insurer pays a fixed amount for a given CT examination, then, regardless of how many CT sections are obtained, the cost to the insurer is the same. However, the cost from the perspective of the radiology group will vary depending on their choice of imaging protocol. Since the cost depends on the perspective, it is possible for the same problem to produce different conclusions, depending on the perspective chosen. Despite the effect of the perspective, investigators in a recent study of economic evaluations of diagnostic tests (30) found that only 5% of the publications reviewed stated the perspective.

Effectiveness Measure

Selection of an appropriate measure of effectiveness requires careful consideration of the relative advantages and disadvantages of each strategy. The outcome, or effectiveness measure, such as the number of cancer cases detected, willingness to pay, life-years saved, or QALYs, should be clearly stated. The U.S. Public Health Service Panel recommends using the QALY, which was arguably the most popular effectiveness measure at the time this article was written (31). This measure incorporates both length and quality of life. Grann et al (32) examined the benefit of prophylactic mastectomy and oophorectomy in 30-year-old women who tested positive for a genetic predisposition to breast cancer. Although prophylactic surgery substantially increased length of life, the entire benefit disappeared when quality of life was included in the analysis. Another benefit of using QALYs is that it is not disease specific, which allows for comparison across studies and diseases. For example, the value of cervical spine screening CT in patients with trauma can be compared with that of MR imaging and core-needle biopsy in the
work-up of suspicious breast lesions in women (33,34).

Patient perception of the value of a health care intervention—with its potential risks and benefits—are part of the definition of any utility measure, including QALY. In some cases this may be problematic, since these values have not been well studied in certain populations, such as children and racial and cultural subgroups. Sometimes there is insufficient data to make the link to long-term outcomes, such as life-years saved. Radiology CEA often use intermediate outcome measures such as correct diagnosis, length of hospital stay, number of emergency department visits saved, willingness to pay, and other resource use measures (21,35).

**BUILDING THE CEA MODEL**

Once the problem has been defined, the next step is to build the model. This is done by creating a decision tree, in which the structure of the model is designed. Afterward, probability, utility, and cost data must be supplied. The credibility of the model will be judged primarily on three factors: the degree to which the model reflects clinical practice, the strength of the data, and the strength or weakness of the assumptions that were made.

**Decision Trees**

A CEA model is based on a decision tree. This tree is a visual representation of the research question(s) and is a powerful analytic tool for complex medical decisions. The decision tree might be described as a horizontal flow chart that depicts all the choices, events, and outcomes. The result is either a positive or negative result. In each of these cases, the result may not be correct. A positive test result usually indicates acute appendicitis but could also be a false-positive result. In a similar way, a negative test result is sometimes a missed diagnosis. If the outcome of interest is a correct diagnosis, then we have reached the end of the tree and designate this with a terminal node. Each outcome is associated with a “pay-off,” or the value associated with that particular outcome. The clinical examination strategy has a similar sequence of events. The result is either a positive or negative clinical examination result. Again, the diagnosis may not be correct. To solve this problem, every possibility following a chance node must be assigned a probability, and each path through the tree must be assigned an associated cost and outcome score.

**Probabilities**

Once the decision tree is constructed, the model parameters must be estimated. Every possibility identified in the decision tree has an associated probability of that event occurring. A patient undergoing imaging has a certain probability of a positive test result. This probability may be estimated from historical data and is usually obtained from the literature. If literature is lacking, data from a pilot study or expert opinion may be substituted. Expert opinion may represent the judgment of a single researcher who is involved in the project or may be based on the consensus of a group of experts, which is obtained through a Delphi panel or survey (37). Although expert opinion is inevitably subject to bias, it may be the only source of a probability for a given event. For example, although there is a substantial body of literature on the use of MR cholangiopancreatography in adults, less is known about its use in children. Expert opinion might be needed to estimate the sensitivity and specificity of MR cholangiopancreatography in children.

**Utility**

Not all health conditions are equal. Utility is a quantity that measures a decision maker’s preference for a particular outcome or state of health. The term “utility” is commonly used interchangeably with the word “value.” Assigning utilities to health conditions is a means of incorporating patient values. To calculate QALYs, it is necessary to assign a utility to every health condition in the model. For example, in a study of prostate screening, Cantor et al (38) required the utility not only of prostate cancer but also of treatment side effects such as impotence and urethral stricture. There are

---

**Figure 1.** Simplified decision tree for a patient with acute abdominal pain at presentation to the emergency department. □ = decision node, ○ = chance node, ◼ = terminal node.

---

**Table 1.** Probabilities for the decision tree for a patient with acute abdominal pain at presentation to the emergency department.
two anchors for utility scores: Dead is 0, and full health is 1. It is possible but rare to have states of health worse than death, which would receive a negative preference weight. Table 2 provides a small sample of published utility values for different conditions.

There is no clearly accepted standard for determining utility scores. The three most common methods are the standard gamble, the time trade-off, and the rating scale, which is often called a visual analog scale (42). The simplest is a rating scale, in which a person rates a state of health by marking a spot on a linear scale, with death and full health marked to represent zero and one. The standard gamble presents a choice between a given outcome, the health state of interest, and a risky gamble, which results in either cure or death. The greater the risk of death the person is willing to take, the lower the utility of the health state. The time trade-off presents a choice between two given outcomes. The first choice is remaining in the specified health state for a specified number of years. The person is then asked how many years of life lived in full health would be equivalent to the first choice. The fewer the years of healthy life the person is willing to accept in trade, the lower the utility of the health state. Experience has shown that standard gamble scales tend to produce the highest scores, with time trade-offs usually being slightly lower and rating scale values being the lowest (43). The Health Utilities Index and EuroQOL-5D are alternative measures that are becoming popular because of their easy applicability (42). These tools consist of a small set of simple questions to patients, and the responses are then mapped automatically to a utility score.

### Costs

Another challenge in CEA is to identify all relevant costs. Costs should not be confused with charges (44): rather, a cost is a valuation of resources expended. A single CEA might include the costs of a diagnostic test, the test interpretation fee, the nursing fee, and the treatment of any complications, as well as patient transportation and parking. When a patient with cancer receives a course of radiation therapy, the analyst should consider the value of the patient’s time and the cost of transportation in addition to the medical cost associated with the therapy. If the patient needs to be accompanied by a caregiver, then the caregiver’s time must also be valued. Whether it is a paid caregiver or a family member, the use of a person’s time is an expenditure of society’s resources. Those resources could otherwise have been used for a different purpose, including leisure time.

Although the terms “direct costs” and “indirect costs” are commonly used in the literature, the terms mean different things to different people, which creates confusion. Some define a direct cost as the value of a resource that is consumed in the provision of an intervention or for any consequences of the intervention (45). In accordance with this definition, the time a patient spends receiving radiation therapy for cancer is considered to be a direct cost because it is an integral part of the treatment. Others define direct costs more narrowly so that they include only the value of resources that are actually paid for (46). With this definition, patient time spent receiving radiation therapy would not be considered a direct cost but rather an indirect cost. With both definitions, the cost due to lost work or leisure time that results from morbidity is an indirect cost. Study authors should explicitly state what costs are included and avoid the use of ambiguous labels.

Depending on which perspective the analysis has assumed, certain costs may be excluded. If the societal perspective is adopted, then all costs must be considered, including lost productivity due to missed work. If the perspective is that of the hospital, costs such as patient and caregiver time would be excluded, since they are not part of the hospital’s financial responsibility. Authors of published analyses often state that they are considering only medical costs. Although this simplifies the analysis, since nonmedical costs are often harder to estimate, it is a potential source of bias. For example, studies have shown that for a noninstitutionalized patient with dementia, the cost of unpaid caregiver services exceeds the medical cost (47, 48).

Perhaps the most accurate method of cost estimation is that known as micro-costing, in which every resource use is identified and quantified into a unit cost (49). It is common to use less labor-intensive methods and substitute proxies for cost, such as Medicare or Medicaid reimbursement. This method has the advantage of using a nationally relevant estimate as opposed to a single facility’s cost. Medicare or Medicaid reimbursement values are public information, and the results may be more generalizable to a reader’s practice.

Another popular technique is to start with a hospital’s charges and then multiply by an adjustment called the cost-to-charge ratio. This ratio can be estimated from the Medicare cost report and is also public information. Although convenient, the cost-to-charge ratio is usually available only for a hospital or department and not for a specific procedure or diagnosis.

### Current Cost Estimates

Cost estimates come from different sources. They either are based on current costs or use cost estimates from the literature. This often results in cost estimates from different years. All cost estimates in a model must be based on a common starting date. If the starting date is the year 2000, then the study authors must explicitly state that all costs are in year 2000 U.S. dollars, and all costs must be adjusted to reflect an estimate of costs in the year 2000. This adjustment is usually made by using government estimates of inflation, such as the medical services component of the consumer price index, which is available on the Web site for the Bureau of Labor Statistics (50). For example, if the inflation rate is estimated at 4%, then a cost estimate of $100 from 1 year ago will now be 4% higher, which results in a revised cost estimate of $100 × 1.04 = $104.

### Time Preference for Money

A CEA model may have a time horizon that spans multiple years into the future. Even without inflation, most people would prefer receiving $1,000 today to receiving it in 20 years. A delay in receipt of the money results in an opportunity cost, since the opportunity to benefit from the use of the money would be lost.

---

**TABLE 2 Examples of Published Utility Values**

<table>
<thead>
<tr>
<th>Condition*</th>
<th>Utility Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full health</td>
<td>1.00</td>
</tr>
<tr>
<td>Side effects from interferon (39)</td>
<td>0.93</td>
</tr>
<tr>
<td>Mild angina (24)</td>
<td>0.90</td>
</tr>
<tr>
<td>Prophylactic mastectomy and oophorectomy (32)</td>
<td>0.86</td>
</tr>
<tr>
<td>Moderate angina (24)</td>
<td>0.70</td>
</tr>
<tr>
<td>Hospitalization for tuberculosis (40)</td>
<td>0.60</td>
</tr>
<tr>
<td>Severe angina (24)</td>
<td>0.50</td>
</tr>
<tr>
<td>Moderate to severe stroke (41)</td>
<td>0.39</td>
</tr>
<tr>
<td>Recurrent stroke (41)</td>
<td>0.12</td>
</tr>
<tr>
<td>Death</td>
<td>0.00</td>
</tr>
</tbody>
</table>

* Numbers in parentheses are reference numbers.
To incorporate the opportunity cost, all future costs are adjusted downward to an equivalent cost in current dollars. The process of devaluing money for each year in the future is called discounting. The calculation is handled in much the same way that old cost estimates are adjusted upward for inflation to create a current cost estimate. However, with discounting, we reverse the process and move backward in time, which causes future costs to become smaller in the present. A medical procedure that costs $100 1 year from now at a discount rate of 3% is converted to a current cost estimate of $100/1.03 = $97.87.

The U.S. Public Health Service panel recommends a 3% discount rate for the baseline analysis (51). Since numerous published studies have used a 5% discount rate, they also recommend analyzing the model by using a 5% discount rate to facilitate comparison with these studies.

**Time Preference for Health**

Although most people find it easy to accept the notion of a time preference for money, the concept is harder to grasp when it is applied to health benefits. Imagine that the government has a choice between implementing a one-time-only health program now versus in 50 years. Assume that the benefits would be the same whenever it is implemented, although it can be implemented only once. Most people would choose to implement the program now.

Although controversial, the notion of discounting health benefits is widely accepted (52,53). Typically, we discount health at the same rate at which we discount cost. If the discount rates differ after adjusting for inflation, the relative relationship between health and money would change over time (53). Although it may be informative to examine what happens if health benefits are not discounted, this approach should be reserved for sensitivity analysis, which is addressed in the next section.

**Assumptions**

CEAs usually require multiple assumptions, which should mirror those made in actual practice. For example, to simplify the model, a CEA may assume that patients undergoing barium enema examination will never have a reaction to contrast material. However, most assumptions are simply due to limitations imposed by the available data. Authors must explicitly state all assumptions and discuss their potential effects so that readers may make their own judgment of the credibility of the CEA.

**ANALYSIS OF THE CEA MODEL**

Once the problem definition and model building are complete, the analysis is performed. The necessary steps in model analysis include the baseline analysis of cost and effectiveness, the incremental CEA, and sensitivity analysis. These analyses use the mean cost and mean effectiveness for each strategy. Each path in the decision tree represents a sequence of events with an associated cost and effectiveness value. These values can be averaged on the basis of the likelihood, or probability, of each path in the tree. For a detailed description of how to analyze decision trees, see Weinstein et al (54).

**Baseline Analysis**

In CEA, cost and effectiveness are examined simultaneously. Most CEAs deal with a “cost-conscious” environment, in which there is no fixed budget but a desire to make fiscally responsible decisions. Incremental CEA is designed to answer the question “Is it worth the extra money to implement strategy B instead of strategy A?” A common approach is to examine a currently accepted standard of care versus a new, more expensive strategy that improves outcomes. When the analysis of the mean cost and effectiveness for each strategy is completed, there are four possible results: (a) Strategy B is less expensive and more effective than strategy A, (b) strategy B is more expensive and less effective than strategy A, (c) strategy B is more expensive and more effective than strategy A, and (d) strategy B is less expensive and less effective than strategy A.

Cases 1 and 2 illustrate the concept of dominance. A strategy dominates another strategy if it is both more effective and less costly. In this situation, the conclusion is obvious, since one strategy is superior in all measures. In cases 3 and 4, trade-offs occur (ie, the more effective strategy comes at the price of additional cost). In these two cases, strategies are compared by using the incremental cost-effectiveness ratio (ICER), which is the ratio of the increase in cost, or incremental cost, to the gain in health benefits, or incremental effectiveness. Incremental CEA results for a hypothetical example are given in Table 3.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Mean Cost ($)</th>
<th>Incremental Cost ($)</th>
<th>Mean Effectiveness (QALYs)</th>
<th>Incremental Effectiveness (QALYs)</th>
<th>ICER ($/QALY)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1,000</td>
<td></td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>11,000</td>
<td>10,000</td>
<td>11</td>
<td>1</td>
<td>10,000</td>
</tr>
</tbody>
</table>

On average, strategy A yields 10 QALYs at a mean cost of $1,000, and strategy B yields a mean of 11 QALYs at a cost of $11,000. In Table 3, the ICER is the incremental cost, $10,000, divided by the incremental effectiveness, 1 QALY, which yields $10,000 per QALY. The question then becomes “Is $10,000 a reasonable price to pay to gain 1 additional QALY?” There is no definite answer. Some have suggested as a rough guideline the following ranges: Less than $20,000 per QALY is clearly cost-effective, greater than $100,000 per QALY is clearly not cost-effective, and in between is possibly cost-effective (55). Others have gone further, adopting $50,000 per QALY as a benchmark (22,56). In the absence of a clear standard, it may be helpful to compare a cost-effectiveness ratio with interventions that society has determined to be worthwhile. Tengs et al (57) compared the cost-effectiveness of 500 lifesaving interventions. A review of this list may provide a sense of relative value, but the reader must bear in mind that society’s decision to adopt a health intervention may have little to do with its cost-effectiveness.

Investigators in some studies also report the mean cost-effectiveness ratio. These ratios can be misleading. Comparing strategies on the basis of mean cost-effectiveness ratios implies that the goal is to pay the lowest amount per unit of effectiveness without regard to the total quantity of health benefits provided. By using mean cost-effectiveness ratios, a strategy that yields 1 QALY for $1 (mean cost-effectiveness ratio = $1/QALY) would be preferred to a strategy that yields 30 QALYs for $60 (mean cost-effectiveness ratio = $2 per QALY). However, it would clearly be preferable to pay the additional
$59 to implement the latter strategy and obtain an additional 29 QALYs, an ICER of just $2.03 per QALY. Incremental CEA assumes that the goal is to maximize health benefits as long as the price of the additional health is reasonable.

CEAs may evaluate more than two strategies. In these cases, each strategy is compared with the next most effective strategy. Since a strategy that gets compared with a particularly poor strategy would have an unfair advantage, it is necessary to remove from analysis any strategy that is both more costly and less effective than another strategy. In a similar way, any strategy with an ICER worse, or higher, than that of a more effective strategy should also be eliminated. The ICERs are then recalculated for each remaining strategy. The recommended strategy is the most effective strategy in which incremental benefits are obtained for a reasonable incremental cost.

Example

Let us return to the simplified decision tree in Figure 1. In this case, we have a patient with abdominal pain at presentation to the emergency department. Two strategies were considered: CT and clinical examination. The outcome measure will be whether the diagnosis is correct, which is coded as 1 if correct or as 0 if incorrect. For simplicity, we will consider only the cost of emergency department evaluation and imaging. We estimated a 25% prevalence of acute appendicitis and 96% sensitivity and specificity for CT. We used estimates of 90% and 85% for the sensitivity and specificity of clinical examination, respectively. These numbers were manipulated to produce the probability estimates that were needed for the tree. The cost of an emergency department evaluation, changing the sensitivity and specificity of clinical examination, respectively. These numbers were manipulated to produce the probability estimates that were needed for the tree.

The convention is to put probabilities under the associated branch. To the right of each terminal node is the payoff. For CEA, the convention is to show the cost first, followed by the effectiveness score, with a slash separating the two. In accordance with the tree, if CT is performed, the test result is positive, and the diagnosis is correct, we see a payoff of 325/1, which means that the cost associated with this path is $325, and the outcome score is 1. The cost was the sum of an emergency department evaluation and the cost of the CT examination ($75 + $250 = $325). The effectiveness score is 1, which designates a correct diagnosis.

All paths associated with clinical examination have a cost of $75, and all paths with an incorrect diagnosis have an outcome score of 0. When the probabilities of each path are calculated, the mean cost and mean effectiveness can be computed by weighting each outcome by its likelihood of occurring. By using these numbers, a cost-effectiveness results table (Table 4) can be created.

We can see that the clinical examination costs a mean of $75 and leads to a correct diagnosis in 86% of all cases. CT costs $325 and leads to a correct diagnosis in 96% of all cases. Using incremental analysis, we find that CT cost an extra $250 and led to a correct diagnosis in an additional 10% of cases. Thus, the ICER of CT, as compared with clinical examination, is $250/0.10 = $2,500 per additional correct diagnosis. In this case, CT would be considered cost-effective, as compared with clinical examination, if $2,500 is considered a reasonable price to pay for a correct diagnosis. This example was simplified for ease of exposition. A proper analysis of this problem would have to consider the consequences that result from false-positive and false-negative examinations, as well as the costs associated with surgery and follow-up care.

Sensitivity Analysis

A model is only as good as the underlying data and assumptions. However, not every parameter or assumption will have a significant effect on the model’s conclusion. It is important to test the robustness of the model with regard to its assumptions and parameter uncertainty. Sensitivity analysis involves changing the parameter values (probabilities, utilities, and costs) used in the model to determine the effect on the conclusion. In the earlier example of imaging for suspected appendicitis in the emergency department, changing the sensitivity of CT from the baseline value of 98% to 90% would evaluate the robustness of the conclusion. Each CEA parameter needs to be evaluated for the range of clinically plausible values, just as randomized con-
trolled trials typically cite CIs in the results. If the decision is robust with regard to changes in a parameter, this strengthens the conclusion of the study. If, however, small changes in a parameter’s value can change the conclusion, then the conclusion is said to be sensitive to that parameter, and the conclusion is less clear.

One-way sensitivity analysis involves changing a single parameter while holding all other parameters at their baseline value. It is common practice to conduct one-way sensitivity analyses on all parameters in the model. Two-way sensitivity analysis looks at the effects of changing two parameters simultaneously. Typically, two-way sensitivity analyses are performed for the most sensitive parameters in the one-way sensitivity analyses, or a pair might be selected because there is an important relationship to be demonstrated.

Sometimes models are constructed for problems for which there are few data. The sensitivity analysis may then be more important than the conclusion. By examining each parameter that is not well known and determining which ones have the greatest effect on the conclusion, it will become clear which parameters need to be examined further. Sensitivity analysis can aid in identifying research priorities. In the CEA study of cervical spine screening CT for trauma, Blackmore et al (33) suggested the need for further research on the probability of developing a severe neurologic deficit from an undiagnosed cervical spine fracture; this was the area of greatest uncertainty in their model.

In an evaluation of the cost-effectiveness of MR imaging and core-needle biopsy of suspicious breast lesions, Hrung et al (34) found that these preoperative tests were cost-effective and strongly influenced by patient preferences. On the basis of the sensitivity analyses, they concluded that more research was needed to determine more precise estimates of MR imaging and core-needle biopsy test performance characteristics.

BARRIERS TO THE USE OF COST-EFFECTIVENESS RESULTS

Several barriers limit the effect of CEAs. Readers of a CEA study may disagree with the set of strategies selected for comparison. Study authors may compare a strategy with one that is not the standard of care, possibly because the standard changed between the time the study began and its publication. Another problem may be that the set of choices considered does not reflect the entire array of clinically reasonable strategies. A reader may dismiss the study as irrelevant because the strategy favored by that individual was not included. As an alternative, if an unappealing strategy is included, credibility may be lost, and the audience may dismiss the entire study.

For a CEA to have an effect, the analysis must be relevant to the decision maker, be it a clinician or a managed care insurer. The study population, or reference case, must be similar to the population treated at the decision maker’s facility. In addition, the time horizon must be relevant. Managed care organizations may give higher priority to studies in which shorter time horizons are used, since their membership usually turns over rapidly.

Even when a CEA has been well designed and is relevant, the results may be viewed skeptically because they are frequently funded by pharmaceutical companies or, less frequently, imaging equipment manufacturers. These companies have the greatest motivation to fund such studies. An economic evaluation may demonstrate that the additional cost of a new drug or technology is reasonable because of the improved health benefit that results. Nonetheless, readers may be skeptical of the vested interest, fearing that there will be bias in the analysis.

Another barrier to more widespread acceptance of CEAs involves timing. If a study appears too early, when few data are available, readers may feel that the data are insufficient and dismiss the study. On the other hand, if a model appears after there are sufficient data to comfortably conduct the CEA, health care providers may have already made their decisions and consider the CEA obsolete.

CONCLUSION

CEA is an objective systematic method for comparing health care strategies on the basis of both cost and effectiveness simultaneously. It is an established research technique for informing health care policymakers, guiding researchers in new directions, and testing strategies that would be difficult to test in actual practice. These studies must be conducted with methodologic rigor. This places a substantial burden on the reader to understand the terminology and methods to properly evaluate published CEAs. The reader is encouraged to review Appendix 2, which provides a recommended reading list of books and articles on the methods, evaluation, and application of CEA in radiology.

There is no doubt that we will see more CEAs in the radiology literature in the future. At the 1998 Workshop on Health Services and Outcomes Research in Radiology, Alvin I. Mushlin, MD, ScM, a distinguished researcher in health outcomes, suggested that “as the importance of diagnostic tests and procedures and their relationship to medical care resources become better recognized, there will be even more interest in economic evaluations, which start with and focus on diagnostic testing. This will create opportunities for new insights into the reasons for the high cost of medical care and will undoubtedly uncover ways to improve quality rationally and decrease waste and unnecessary expenditures” (58). It is essential that the readers of this literature become familiar with the basic vocabulary, rationale, and standard methods of CEA. By improving our knowledge and understanding of this state-of-the-art research tool, the radiology community will have a greater ability to participate in health care policy setting and decision-making locally and nationally.

APPENDIX 1: CHECKLIST OF THE COMPONENTS OF A CEA

I. Problem Definition: Is the question well defined? Are all strategies considered?
   a. Reference case
   b. Strategies
   c. Time horizon
   d. Perspective
   e. Effectiveness measure(s)
II. Model Building: Are all important relevant consequences and costs included?
   a. Decision tree structure
   b. Probability, utility, and cost data sources
   c. Time preference for money and health
   d. Assumptions
III. Model Analysis: Is incremental analysis performed? Is uncertainty allowed for?
   a. Baseline cost and effectiveness
   b. Incremental CEA
   c. Sensitivity analysis
**APPENDIX 2: RECOMMENDED READINGS**

**Methods of CEA**


**Examples of Published CEAs**


How to Evaluate CEAs


Acknowledgments: We thank J. Shannon Swan, MD, Scott B. Cantor, PhD, and Carlos Sivit, MD, for their thoughtful suggestions that enhanced the readability and clarity of this manuscript.

References


10. Sivit, MD, for their thoughtful suggestions that enhanced the readability and clarity of this manuscript.


